

# Recovery of signals under the high order RIP condition via prior support information

Wengu Chen and Yaling Li <sup>‡</sup>

March 14, 2016

## Abstract

In this paper we study the recovery conditions of weighted  $l_1$  minimization for signal reconstruction from incomplete linear measurements when partial prior support information is available. We obtain that a high order RIP condition can guarantee stable and robust recovery of signals in bounded  $l_2$  and Dantzig selector noise settings. Meanwhile, we not only prove that the sufficient recovery condition of weighted  $l_1$  minimization method is weaker than that of standard  $l_1$  minimization method, but also prove that weighted  $l_1$  minimization method provides better upper bounds on the reconstruction error in terms of the measurement noise and the compressibility of the signal, provided that the accuracy of prior support estimate is at least 50%. Furthermore, the condition is proved sharp.

Keywords: Compressed sensing, restricted isometry property, weighted  $l_1$  minimization.

## 1 Introduction

Compressed sensing is a new type of sampling theory that admits that high dimensional sparse signals can be reconstructed through fewer measurements than their ambient dimension. The central goal in compressed sensing is to recover a signal  $x \in \mathbb{R}^N$  based on  $A$  and  $y$  from the following model:

$$y = Ax + z \tag{1}$$

---

<sup>\*</sup>W. Chen is with Institute of Applied Physics and Computational Mathematics, Beijing, 100088, China, e-mail: chenwg@iapcm.ac.cn.

<sup>†</sup>Y. Li is with Graduate School, China Academy of Engineering Physics, Beijing, 100088, China, e-mail: leeyaling@126.com.

<sup>‡</sup>This work was supported by the NSF of China (Nos.11271050, 11371183) and Beijing Center for Mathematics and Information Interdisciplinary Sciences (BCMIIS).

where sensing matrix  $A \in \mathbb{R}^{n \times N}$  with  $n \ll N$ , i.e., using very few measurements,  $y \in \mathbb{R}^n$  is a vector of measurements, and  $z \in \mathbb{R}^n$  is the measurement error. In past decade, compressed sensing has triggered considerable research in a number of fields including applied mathematics, statistics, electrical engineering, seismology and signal processing. Compressed sensing is especially promising in applications where taking measurements is costly, e.g., hyperspectral imaging [11], as well as in applications where the ambient dimension of the signal is very large, e.g., medical image reconstruction [18], DNA microarrays [20], radar system [1, 13, 23].

To reconstruct the signal  $x$  from (1), Candès and Tao [10] proposed the following constrained  $l_1$  minimization method:

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \|x\|_1 \quad \text{subject to} \quad \|y - Ax\|_2 \leq \epsilon. \quad (2)$$

It is well known that  $l_1$  minimization is a convex relaxation of  $l_0$  minimization and is polynomial-time solvable. And it has been shown that  $l_1$  minimization is an effective way to recover sparse signals in many settings [3–9, 19]. Cai and Zhang [6, 7] established sharp restricted isometry conditions to achieve the exact and stable recovery of signals in both noiseless and noisy cases via  $l_1$  minimization method.

Note that compressed sensing is a nonadaptive data acquisition technique and  $l_1$  minimization method (2) is itself nonadaptive because no prior information on the signal  $x$  is used. In practical examples, however, the estimate of the support of the signal or of its largest coefficients may be possible to be drawn. For example, support estimation of the previous time instant may be applied to recover time sequences of sparse signals iteratively. Incorporating prior information is very useful for recovering signals from compressive measurements. Thus, the following weighted  $l_1$  minimization method which incorporates partial support information of the signals has been introduced to replace standard  $l_1$  minimization

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \|x\|_{1,w} \quad \text{subject to} \quad \|y - Ax\|_2 \leq \epsilon, \quad (3)$$

where  $w \in [0, 1]^N$  and  $\|x\|_{1,w} = \sum_i w_i |x_i|$ . Reconstructing compressively sampled signals with partially known support has been previously studied in the literature; see [2, 12, 14–17, 21]. Borries, Miosso and Potes in [2], Khajehnejad *et al.* in [15], and Vaswani and Lu in [21] introduced the problem of signal recovery with partially known support independently. The works by Borries *et al.* in [2], Vaswani and Lu in [16, 21, 22] and Jacques in [14] incorporated known support information using weighted  $l_1$  minimization approach with zero weights on the known support, namely, given a support estimate  $\tilde{T} \subset \{1, 2, \dots, N\}$  of unknown signal  $x$ , setting  $w_i = 0$  whenever  $i \in \tilde{T}$  and  $w_i = 1$  otherwise, and derived sufficient recovery conditions. Friedlander *et al.* in [12] extended weighted  $l_1$  minimization approach to nonzero weights. They allow the weights  $w_i = \omega \in [0, 1]$  if  $i \in \tilde{T}$ . Since Friedlander *et al.* incorporated the prior support information

and consider the accuracy of the support estimate, they derived the stable and robust recovery guarantees for weighted  $l_1$  minimization which generalize the results of Candès, Romberg and Tao in [9] stated below. Friedlander *et al.* [12] pointed out that once at least 50% of the support information is accurate, the weighted  $l_1$  minimization method (3) can stably and robustly recover any signals under weaker sufficient conditions than the analogous conditions for standard  $l_1$  minimization method (2). In addition, the weighted  $l_1$  minimization method (3) gives better upper bounds on the reconstruction error. Furthermore, they also pointed out sufficient conditions are weaker than those of [21] when  $\omega = 0$ .

To recover sparse signals via constrained  $l_1$  minimization, Candès and Tao [10] introduced the notion of Restricted Isometry Property (RIP), which is one of the most commonly used frameworks for compressive sensing. The definition of RIP is as follows.

**Definition 1.1.** Let  $A \in \mathbb{R}^{n \times N}$  be a matrix and  $1 \leq k \leq N$  is an integer. The restricted isometry constant (RIC)  $\delta_k$  of order  $k$  is defined as the smallest nonnegative constant that satisfies

$$(1 - \delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k)\|x\|_2^2,$$

for all  $k$ -sparse vectors  $x \in \mathbb{R}^N$ . A vector  $x \in \mathbb{R}^N$  is  $k$ -sparse if  $|\text{supp}(x)| \leq k$ , where  $\text{supp}(x) = \{i : x_i \neq 0\}$  is the support of  $x$ . When  $k$  is not an integer, we define  $\delta_k$  as  $\delta_{\lceil k \rceil}$ , where  $\lceil k \rceil$  denotes the smallest integer strictly bigger than  $k$ .

Candès, Romberg and Tao [9] showed that the condition  $\delta_{ak} + a\delta_{(a+1)k} < a - 1$  with  $a \in \frac{1}{k}\mathbb{Z}$  and  $a > 1$  is sufficient for stable and robust recovery of all signals using  $l_1$  minimization method (2). Cai and Zhang [6] improved the result of Candès, Romberg and Tao [9] and proved that the condition  $\delta_{tk} < \sqrt{\frac{t-1}{t}}$  with  $t \geq 4/3$  can guarantee the exact recovery of all  $k$ -sparse signals in the noiseless case and stable recovery of approximately sparse signals in the noise case by  $l_1$  minimization method (2). Furthermore, Cai and Zhang proved that for any  $\epsilon > 0$ ,  $\delta_{tk} < \sqrt{\frac{t-1}{t}} + \epsilon$  fails to ensure the exact reconstruction of all  $k$ -sparse signals and stable reconstruction of approximately sparse signals for large  $k$ .

In [6], Cai and Zhang use the following  $l_1$  minimization

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \|x\|_1 \quad \text{subject to} \quad \|y - Ax\|_2 \in \mathcal{B}, \quad (4)$$

where  $\mathcal{B}$  is a bounded set determined by the noise structure, and  $\mathcal{B}$  is especially taken to be  $\{0\}$  in the noiseless case. They consider two types of noise settings

$$\mathcal{B}^{l_2}(\epsilon) = \{z : \|z\|_2 \leq \epsilon\} \quad (5)$$

and

$$\mathcal{B}^{DS}(\epsilon) = \{z : \|A^T z\|_\infty \leq \epsilon\}. \quad (6)$$

In this paper, we adopt the corresponding weighted  $l_1$  minimization method:

$$\begin{aligned} & \underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \|x\|_{1,w} \quad \text{subject to} \quad y - Ax \in \mathcal{B} \\ & \text{with} \quad w_i = \begin{cases} 1, & i \in \tilde{T}^c \\ \omega, & i \in \tilde{T}. \end{cases} \end{aligned} \quad (7)$$

where  $0 \leq \omega \leq 1$  and  $\tilde{T} \subset \{1, 2, \dots, N\}$  a given support estimate of unknown signal  $x$ .  $\mathcal{B}$  is also a bounded set determined by the noise settings (5) and (6). Our goal is to generalize the results of Cai and Zhang [6] via the weighted  $l_1$  minimization method (7). We establish the high order RIP condition for the stable and robust recovery of signals with partially known support information from (1). We also show that the recovery by weighted  $l_1$  minimization method (7) is stable and robust under weaker sufficient conditions compared to the standard  $l_1$  minimization method (4) when we have the partial support information with accuracy better than 50%.

The rest of the paper is organized as follows. In Section 2, we will introduce some notations and some basic lemmas that will be used. The main results are given in Section 3, and the proofs of our main results are presented in Section 4.

## 2 Preliminaries

Let us begin with basic notations. For arbitrary  $x \in \mathbb{R}^N$ , let  $x_k$  be its best  $k$ -term approximation.  $x_{\max(k)}$  is defined as  $x$  with all but the largest  $k$  entries in absolute value set to zero, and  $x_{-\max(k)} = x - x_{\max(k)}$ . Let  $T_0$  be the support of  $x_k$ , i.e.,  $T_0 = \text{supp}(x_k)$ , with  $T_0 \subseteq \{1, \dots, N\}$  and  $|T_0| \leq k$ . Let  $\tilde{T} \subseteq \{1, \dots, N\}$  be the support estimate of  $x$  with  $|\tilde{T}| = \rho k$ , where  $\rho \geq 0$  represents the ratio of the size of the estimated support to the size of the actual support of  $x_k$  (or the support of  $x$  if  $x$  is  $k$ -sparse). Denote  $\tilde{T}_\alpha = T_0 \cap \tilde{T}$  and  $\tilde{T}_\beta = T_0^c \cap \tilde{T}$  with  $|\tilde{T}_\alpha| = \alpha|\tilde{T}| = \alpha\rho k$  and  $|\tilde{T}_\beta| = \beta|\tilde{T}| = \beta\rho k$ , where  $\alpha$  denotes the ratio of the number of indices in  $T_0$  that were accurately estimated in  $\tilde{T}$  to the size of  $\tilde{T}$  and  $\alpha + \beta = 1$ . For arbitrary nonnegative number  $\xi$ , we denote by  $[[\xi]]$  an integer satisfying  $\xi \leq [[\xi]] < \xi + 1$ .

Cai and Zhang developed a new elementary technique which is a key technical tool for the proof of the main result (see Theorem 3.1). It states that any point in a polytope can be represented as a convex combination of sparse vectors ([6], Lemma 1.1). Another key technical tool for our proof was Lemma 2.2 introduced by Cai and Zhang ([8], Lemma 5.3). The specific contents are presented in Lemmas 2.1 and 2.2, respectively.

**Lemma 2.1** ([6], Lemma 1.1). *For a positive number  $\alpha$  and a positive integer  $k$ , define the polytope  $T(\alpha, k) \subset \mathbb{R}^d$  by*

$$T(\alpha, k) = \{v \in \mathbb{R}^d : \|v\|_\infty \leq \alpha, \|v\|_1 \leq k\alpha\}.$$

For any  $v \in \mathbb{R}^d$ , define the set of sparse vectors  $U(\alpha, k, v) \subset \mathbb{R}^d$  by

$$U(\alpha, k, v) = \{u \in \mathbb{R}^d : \text{supp}(u) \subseteq \text{supp}(v), \|u\|_0 \leq k, \\ \|u\|_1 = \|v\|_1, \|u\|_\infty \leq \alpha\},$$

where  $\|u\|_0 = |\text{supp}(u)|$ . Then any  $v \in T(\alpha, k)$  can be expressed as

$$v = \sum_{i=1}^N \lambda_i u_i,$$

where  $u_i \in U(\alpha, k, v)$  and  $0 \leq \lambda_i \leq 1, \sum_{i=1}^N \lambda_i = 1$ .

**Lemma 2.2** ([8], Lemma 5.3). Assume  $m \geq k, a_1 \geq a_2 \geq \dots \geq a_m \geq 0, \sum_{i=1}^k a_i \geq \sum_{i=k+1}^m a_i$ , then for all  $\alpha \geq 1$ ,

$$\sum_{j=k+1}^m a_j^\alpha \leq \sum_{i=1}^k a_i^\alpha.$$

More generally, assume  $a_1 \geq a_2 \geq \dots \geq a_m \geq 0, \lambda \geq 0$  and  $\sum_{i=1}^k a_i + \lambda \geq \sum_{i=k+1}^m a_i$ , then for all  $\alpha \geq 1$ ,

$$\sum_{j=k+1}^m a_j^\alpha \leq k \left( \sqrt[\alpha]{\frac{\sum_{i=1}^k a_i^\alpha}{k}} + \frac{\lambda}{k} \right)^\alpha.$$

As we mentioned in the introduction, Cai and Zhang [6] provided the sharp sufficient condition to recover sparse signals and approximately sparse signals via  $l_1$  minimization (4). Their main result can be stated as below.

**Theorem 2.1** ([6], Theorem 2.1). Let  $y = Ax + z$  with  $\|z\|_2 \leq \varepsilon$  and  $\hat{x}^{l_2}$  is the minimizer of (4) with  $\mathcal{B} = \mathcal{B}^{l_2}(\eta) = \{z : \|z\|_2 \leq \eta\}$  for some  $\eta \geq \varepsilon$ . If

$$\delta_{tk} < \sqrt{\frac{t-1}{t}} \quad (8)$$

for some  $t \geq 4/3$ , then

$$\|\hat{x}^{l_2} - x\|_2 \leq C_0(\varepsilon + \eta) + C_1 \frac{2\|x_{-\max(k)}\|_1}{\sqrt{k}}, \quad (9)$$

where

$$C_0 = \frac{\sqrt{2t(t-1)(1+\delta_{tk})}}{t(\sqrt{(t-1)/t} - \delta_{tk})}, \\ C_1 = \frac{\sqrt{2}\delta_{tk} + \sqrt{t(\sqrt{(t-1)/t} - \delta_{tk})\delta_{tk}}}{t(\sqrt{(t-1)/t} - \delta_{tk})} + 1. \quad (10)$$

Let  $y = Ax + z$  with  $\|A^T z\|_\infty \leq \varepsilon$  and  $\hat{x}^{DS}$  is the minimizer of (4) with  $\mathcal{B} = \mathcal{B}^{DS}(\eta) = \{z : \|A^T z\|_\infty \leq \eta\}$  for some  $\eta \geq \varepsilon$ . If  $\delta_{tk} < \sqrt{\frac{t-1}{t}}$  for some  $t \geq 4/3$ , then

$$\|\hat{x}^{DS} - x\|_2 \leq C'_0(\varepsilon + \eta) + C'_1 \frac{2\|x_{-\max(k)}\|_1}{\sqrt{k}}, \quad (11)$$

where

$$C'_0 = \frac{\sqrt{2t^2(t-1)k}}{t(\sqrt{(t-1)/t} - \delta_{tk})}, \quad C'_1 = C_1. \quad (12)$$

Note that Theorem 2.1 always holds for  $t > 1$ , and the condition  $t \geq 4/3$  ensures that (8) is sharp.

Friedlander et al. [12] used the prior support information to recover any signals by weighted  $l_1$  minimization (7). The following theorem was showed in [12].

**Theorem 2.2** ([12], Theorem 3). *Let  $x \in \mathbb{R}^N$  be an arbitrary signal and  $y = Ax + z$  with  $\|z\|_2 \leq \varepsilon$ . Define  $x_k$  be its best  $k$ -term approximation with  $\text{supp}\{x_k\} = T_0$ . Let  $\tilde{T} \subseteq \{1, \dots, N\}$  be an arbitrary set and define  $\rho \geq 0$  and  $0 \leq \alpha \leq 1$  such that  $|\tilde{T}| = \rho k$  and  $|\tilde{T} \cap T_0| = \alpha \rho k$ . Suppose that there exists an  $a \in \frac{1}{k}\mathbb{Z}$  with  $a \geq (1 - \alpha)\rho$  and  $a > 1$ . If the measurement matrix  $A$  has RIP satisfying*

$$\delta_{ak} + \frac{a}{\gamma^2} \delta_{(a+1)k} < \frac{a}{\gamma^2} - 1, \quad (13)$$

where  $\gamma = \omega + (1 - \omega)\sqrt{1 + \rho - 2\alpha\rho}$  for some given  $0 \leq \omega \leq 1$ . Then the solution  $\hat{x}$  to (7) with (5) obeys

$$\|\hat{x} - x\|_2 \leq C''_0(2\varepsilon) + C''_1 \frac{2\left(\omega\|x - x_k\|_1 + (1 - \omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1\right)}{\sqrt{k}}, \quad (14)$$

where

$$\begin{aligned} C''_0 &= \frac{1 + \frac{\gamma}{\sqrt{a}}}{\sqrt{1 - \delta_{(a+1)k}} - \frac{\gamma}{\sqrt{a}}\sqrt{1 + \delta_{ak}}}, \\ C''_1 &= \frac{a^{-1/2}(\sqrt{1 - \delta_{(a+1)k}} + \sqrt{1 + \delta_{ak}})}{\sqrt{1 - \delta_{(a+1)k}} - \frac{\gamma}{\sqrt{a}}\sqrt{1 + \delta_{ak}}}. \end{aligned} \quad (15)$$

**Remark 2.1** ([12], Remarks 3.3 and 3.4). *If  $A$  satisfies*

$$\delta_{(a+1)k} < \delta_a^\omega := \frac{a - \gamma^2}{a + \gamma^2}, \quad (16)$$

where  $\gamma = \omega + (1 - \omega)\sqrt{1 + \rho - 2\alpha\rho}$ , then Theorem 2.2 holds with same constants. If

$$\delta_{2k} < \left(\sqrt{2}\gamma + 1\right)^{-1}, \quad (17)$$

then weighted  $l_1$  minimization (7) with (5) can stably and robustly recover the original signal.

### 3 Main results

**Theorem 3.1.** Suppose that  $x \in \mathbb{R}^N$  be an arbitrary signal and  $x_k$  be its best  $k$ -term approximation supported on  $T_0 \subseteq \{1, \dots, N\}$  with  $|T_0| \leq k$ . Let  $\tilde{T} \subseteq \{1, \dots, N\}$  be an arbitrary set and denote  $\rho \geq 0$  and  $0 \leq \alpha \leq 1$  such that  $|\tilde{T}| = \rho k$  and  $|\tilde{T} \cap T_0| = \alpha \rho k$ . Let  $y = Ax + z$  with  $\|z\|_2 \leq \varepsilon$  and  $\hat{x}^{l_2}$  is the minimizer of (7) with (5). If the measurement matrix  $A$  satisfies RIP with

$$\delta_{tk} < \delta_t^\omega := \sqrt{\frac{t-d}{t-d+\gamma^2}} \quad (18)$$

for  $t > d$ , where  $\gamma = \omega + (1-\omega)\sqrt{1+\rho-2\alpha\rho}$  and

$$d = \begin{cases} 1, & \omega = 1 \\ 1 - \alpha\rho + a, & 0 \leq \omega < 1 \end{cases}$$

with  $a = \max\{\alpha, \beta\}\rho$ . Then

$$\|\hat{x}^{l_2} - x\|_2 \leq D_0(2\varepsilon) + D_1 \frac{2\left(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1\right)}{\sqrt{k}}, \quad (19)$$

where

$$\begin{aligned} D_0 &= \frac{\sqrt{2(t-d)(t-d+\gamma^2)(1+\delta_{tk})}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})}, \\ D_1 &= \frac{\sqrt{2}\delta_{tk}\gamma + \sqrt{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})\delta_{tk}}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})} + \frac{1}{\sqrt{d}}. \end{aligned} \quad (20)$$

Let  $y = Ax + z$  with  $\|A^T z\|_\infty \leq \varepsilon$ . Assume that  $\hat{x}^{DS}$  is the minimizer of (7) with (6) and the matrix  $A$  satisfies RIP (18). Then

$$\|\hat{x}^{DS} - x\|_2 \leq D'_0(2\varepsilon) + D'_1 \frac{2\left(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1\right)}{\sqrt{k}}, \quad (21)$$

where

$$D'_0 = \frac{\sqrt{2(t-d)(t-d+\gamma^2)[\lceil tk \rceil]}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})}, \quad D'_1 = D_1. \quad (22)$$

**Remark 3.1.** In Theorem 3.1, every signal  $x \in \mathbb{R}^N$  can be stably and robustly recovered. And if  $\mathcal{B} = \{0\}$  and  $x$  is a  $k$ -sparse vector, then Theorem 3.1 ensures exact recovery of the signal  $x$ .

For Gaussian noise case, the above results on the bounded noise case can be directly applied to yield the corresponding results by using the same argument as in [3, 5]. The concrete content is stated as follows.

**Remark 3.2.** Let  $x \in \mathbb{R}^N$  be an arbitrary signal and  $x_k$  be its best  $k$ -term approximation supported on  $T_0 \subseteq \{1, \dots, N\}$  with  $|T_0| \leq k$ . Let  $\tilde{T} \subseteq \{1, \dots, N\}$  be an arbitrary set and define  $\rho \geq 0$  and  $0 \leq \alpha \leq 1$  such that  $|\tilde{T}| = \rho k$  and  $|\tilde{T} \cap T_0| = \alpha \rho k$ . Assume that  $z \sim \mathcal{N}_n(0, \sigma^2 I)$  in (1) and  $\delta_{tk} < \sqrt{\frac{t-d}{t-d+\gamma^2}}$  for  $t > d$ . Let  $\mathcal{B}^{l_2} = \{z : \|z\|_2 \leq \sigma \sqrt{n + 2\sqrt{n \log n}}\}$  and  $\mathcal{B}^{DS} = \{z : \|A^T z\|_\infty \leq \sigma \sqrt{2 \log N}\}$ .  $\hat{x}^{l_2}$  and  $\hat{x}^{DS}$  are the minimizer of (7) with  $\mathcal{B}^{l_2}$  and  $\mathcal{B}^{DS}$ , respectively. Then, with probability at least  $1 - 1/n$ ,

$$\|\hat{x}^{l_2} - x\|_2 \leq D_0(2\sigma \sqrt{n + 2\sqrt{n \log n}}) + D_1 \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1 - \omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k}},$$

and

$$\|\hat{x}^{DS} - x\|_2 \leq D'_0(2\sigma \sqrt{2 \log N}) + D'_1 \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1 - \omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k}},$$

with probability at least  $1 - 1/\sqrt{\pi \log N}$ . Here  $d = \begin{cases} 1, & \omega = 1 \\ 1 - \alpha \rho + a, & 0 \leq \omega < 1 \end{cases}$  with  $a = \max\{\alpha, \beta\}\rho$ , and  $\gamma = \omega + (1 - \omega)\sqrt{1 + \rho - 2\alpha\rho}$ .

**Theorem 3.2.** Let  $d = 1$  and  $t \geq 1 + \frac{(1 - \sqrt{1 - \gamma^2})^2}{\gamma^2 + 2(1 - \sqrt{1 - \gamma^2})}$ . For any  $\varepsilon > 0$  and  $k \geq \frac{6}{\varepsilon}$ . Then there exists a sensing matrix  $A \in \mathbb{R}^{n \times N}$  with  $\delta_{tk} < \sqrt{\frac{t-d}{t-d+\gamma^2}} + \varepsilon$  and some  $k$ -sparse signal  $x_0$  such that

- (1) In the noiseless case, i.e.,  $y = Ax_0$ , the weighted  $l_1$  minimization (7) can not exactly recover the  $k$ -sparse signal  $x_0$ , i.e.,  $\hat{x} \neq x_0$ , where  $\hat{x}$  is the solution to (7).
- (2) In the noise case, i.e.,  $y = Ax_0 + z$ , for any bounded noise setting  $\mathcal{B}$ , the weighted  $l_1$  minimization (7) can not stably recover the  $k$ -sparse signal  $x_0$ , i.e.,  $\hat{x} \nrightarrow x_0$  as  $z \rightarrow 0$ , where  $\hat{x}$  is the solution to (7).

**Proposition 3.1.** (1) If  $\omega = 1$ , then  $d = 1$  and  $\gamma = 1$ . The sufficient condition (18) of Theorem 3.1 is identical to (8) in Theorem 2.1 and  $D_0 = C_0, D_1 = C_1, D'_0 = C'_0, D'_1 = C'_1$ . Moreover, the condition is sharp if  $t \geq \frac{4}{3}$ .

- (2) If  $\alpha = \frac{1}{2}$ , then  $d = 1, \gamma = 1$ . The sufficient condition (18) of Theorem 3.1 is identical to that of Theorem 2.1 with (8) and  $D_0 = C_0, D_1 = C_1, D'_0 = C'_0, D'_1 = C'_1$ . Moreover, if  $t \geq \frac{4}{3}$ , the condition is sharp.



- (3) Assume that  $0 \leq \omega < 1$ . If  $\alpha > \frac{1}{2}$ , then  $d = 1$  and  $\gamma < 1$ . The sufficient condition (18) in Theorem 3.1 is weaker than (8) in Theorem 2.1 and (16) in Remark 2.1. Then  $D_0 < C_0$ ,  $D_1 < C_1$ ,  $D_0 < C_0''$ ,  $D_1 < C_1''$ . When  $t = 2$ , the sufficient condition (18) in Theorem 3.1 is weaker than (17) in Remark 2.1.

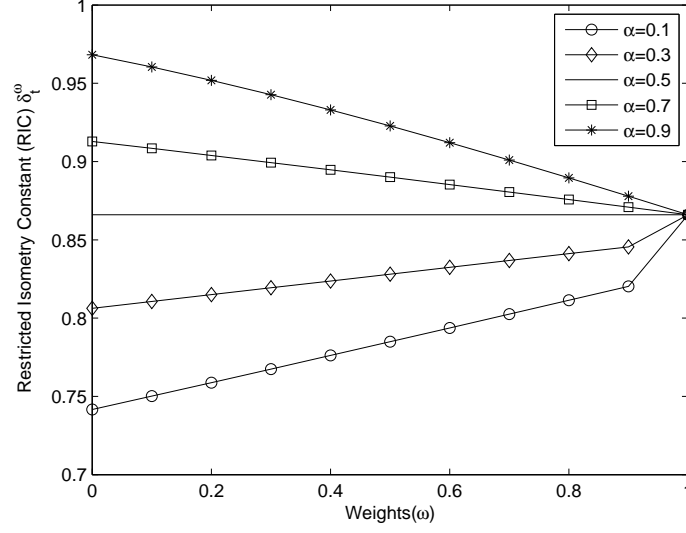
Fig. 1 illustrates how the sufficient conditions on the RIP constants given in (18) and the stability constants given in (20) change with  $\omega$  for different values of  $\alpha$  in the case of weighted  $l_1$  when  $t = 4$ . Note that (18) reduces to (8), and (20) reduces to (10) if  $\omega = 1$  or  $\alpha = 0.5$ . In Fig. 1 (a), we plot  $\delta_t^\omega$  versus  $\omega$  with different values of  $\alpha$  when  $t = 4$ . We observe that the bound on RIP constant gets larger as  $\alpha$  increases. That is to say, the sufficient condition on the RIP constant becomes weaker as  $\alpha$  increases. For example, if 90% of the support estimate is accurate and  $\omega = 0.4$ , we have  $\delta_t^\omega = 0.9330$ , however  $\delta_t^1 = 0.8660$  of standard  $l_1$ . Figs. 1(b) and 1(b') show that the constant  $D_0$  decreases as  $\alpha$  increases with  $\delta_{tk} = 0.1$  and  $\delta_{tk} = 0.6$ , respectively. But Fig. 1(c) demonstrates the constant  $D_1$  with  $\alpha \neq 0.5$  is smaller than that with  $\alpha = 0.5$  when  $\delta_{tk} = 0.1$ . Fig. 1(c') illustrates that the constant  $D_1$  decreases as  $\alpha$  increases with  $\delta_{tk} = 0.6$ . From above recovery results by standard  $l_1$  and weighted  $l_1$ , we see that if the partial support estimate is more than 50% accurate, i.e.  $\alpha > 0.5$ , the measurement matrix  $A$  for signal recovery by weighted  $l_1$  satisfies weaker conditions than the analogous conditions for recovery by standard  $l_1$ . Moreover, we have better upper bounds when  $\alpha > 0.5$  than those of standard  $l_1$ .

Fig. 2 compares the sufficient recovery conditions  $\delta_t^\omega$  in (18) and  $\delta_a^\omega$  in (16) as well as stability constants in (20) and (15) with various  $\alpha$  when  $t = 4$ ,  $a = 3$  and  $\delta_{tk} = \delta_{(a+1)k} = 0.1$ . Here we plot  $\delta_t^\omega$  and  $\delta_a^\omega$  as well as (20) and (15) versus  $\omega$  with various  $\alpha$ . Fig.2(d) illustrates  $\delta_t^\omega$  is larger than  $\delta_a^\omega$  under the same support estimate. Moreover, Figs. 2(e) and 2(f) describe that constants  $D_0$  and  $D_1$  are always smaller than  $C_0''$  and  $C_1''$ , respectively. These results state that the sufficient condition (18) is weaker than (16), and error bound constants (20) in Theorem 3.1 are better than those (15) in Theorem 2.2.

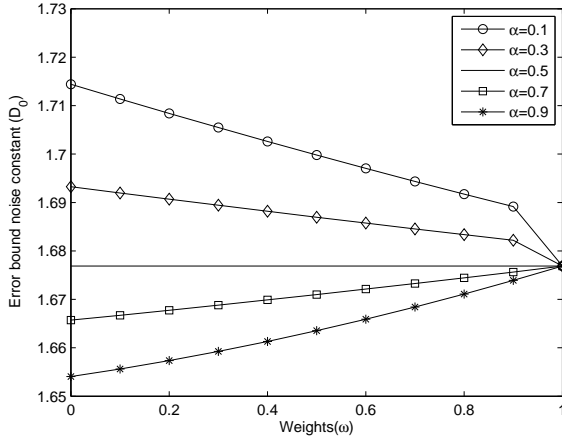
## 4 Proofs

*Proof of Theorem 3.1.* Firstly, we show the estimate (19). Let  $\hat{x}^{l_2} = x + h$ , where  $x$  is the original signal and  $\hat{x}^{l_2}$  is the minimizer of (7) with (5). Now assume that  $tk$  is an integer. We use the following inequality which has been shown by Friedlander *et al.* (see (21) in [12]).

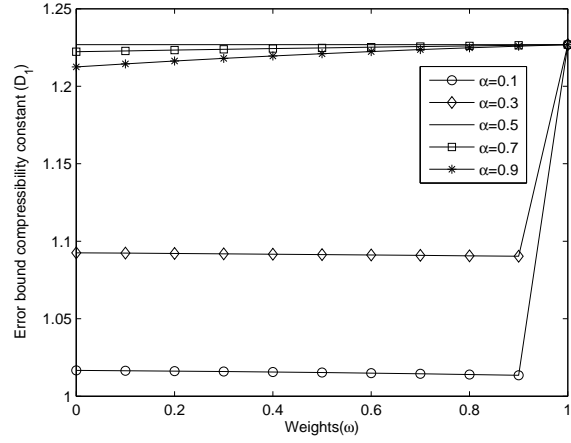
$$\|h_{T_0^c}\|_1 \leq \omega \|h_{T_0}\|_1 + (1 - \omega) \|h_{T_0 \cup \tilde{T} \setminus \tilde{T}_\alpha}\|_1 + 2 \left( \omega \|x_{T_0^c}\|_1 + (1 - \omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right). \quad (23)$$



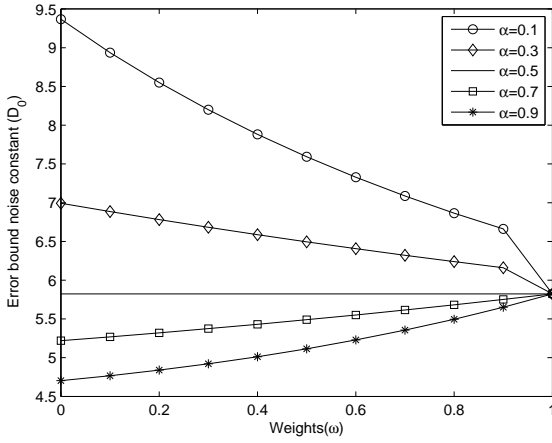
(a)  $\delta_t^\omega$  versus  $\omega$



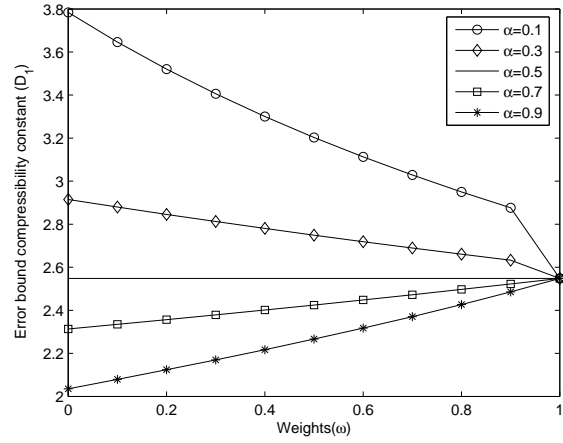
(b)  $D_0$  versus  $\omega$



(c)  $D_1$  versus  $\omega$

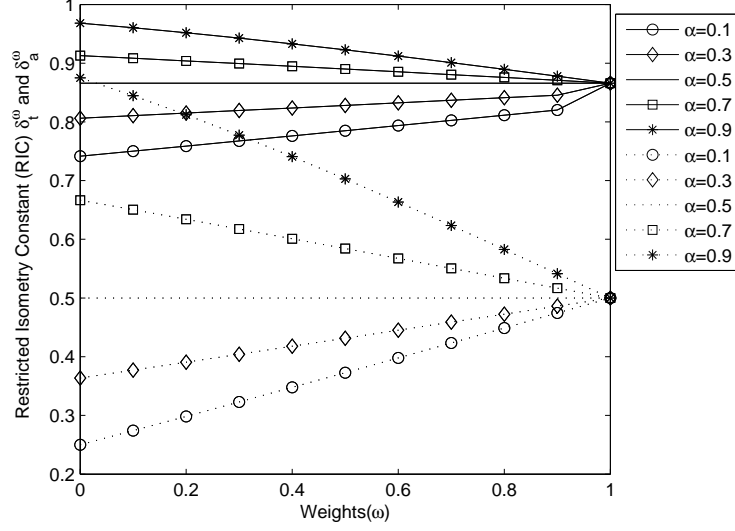


(b')  $D_0$  versus  $\omega$

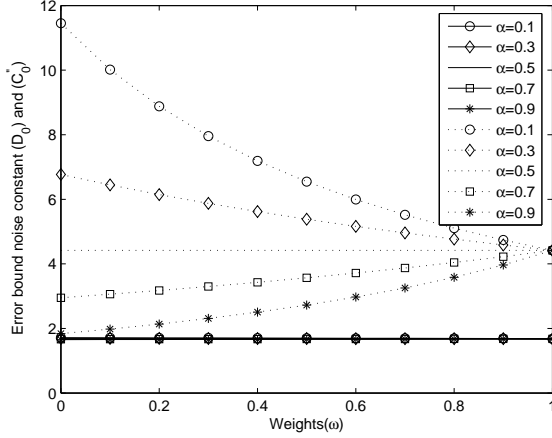


(c')  $D_1$  versus  $\omega$

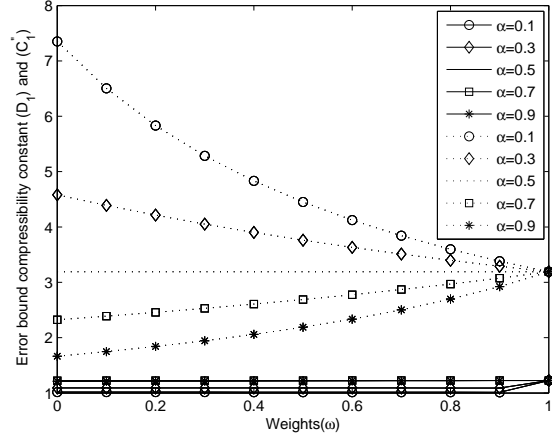
Figure 1: Comparison of the sufficient conditions for recovery and stability constants for weighted  $l_1$  reconstruction with various  $\alpha$ . In all the figures, we set  $t = 4$  and  $\rho = 1$ . In (b) and (c), we fix  $\delta_{tk} = 0.1$ . In (b') and (c'), we fix  $\delta_{tk} = 0.6$ .



(d)  $\delta_t^\omega$  and  $\delta_a^\omega$  versus  $\omega$



(e)  $D_0$  and  $C_0''$  versus  $\omega$



(f)  $D_1$  and  $C_1''$  versus  $\omega$

Figure 2: Comparison between the bounds of sufficient recovery conditions  $\delta_t^\omega$  in (18) and  $\delta_a^\omega$  in (16) as well as stability constants in (20) and (15) with various  $\alpha$ . In all the figures, we set  $t = 4$ ,  $a = 3$  and  $\rho = 1$ . The solid describe our main results and the dotted describe the results of Friedlander et al. [12]. In (e) and (f), we fix  $\delta_{tk} = \delta_{(a+1)k} = 0.1$  and  $\delta_{ak} = 0.05$ .

Let  $\tilde{T}_0 = T_0 \setminus \tilde{T}_\alpha$ , and  $T_1$  indexes the  $ak$  largest in magnitude coefficients of  $h_{\tilde{T}_0^c}$ , where  $|T_1| = ak$  and  $a = \max\{\alpha, \beta\}\rho$ . Denote

$$T_{01} = \begin{cases} T_0, & \omega = 1, \\ \tilde{T}_0 \cup T_1, & 0 \leq \omega < 1. \end{cases}$$

Clearly,  $|T_{01}| = dk$  where

$$d = \begin{cases} 1, & \omega = 1 \\ 1 - \alpha\rho + a, & 0 \leq \omega < 1 \end{cases}.$$

From (23) and  $d \geq 1$ , it is clear that

$$\|h_{-\max(dk)}\|_1 \leq \omega \|h_{T_0}\|_1 + (1 - \omega) \|h_{T_0 \cup \tilde{T} \setminus \tilde{T}_\alpha}\|_1 + 2 \left( \omega \|x_{T_0^c}\|_1 + (1 - \omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right). \quad (24)$$

Let

$$r = \frac{1}{k} \left[ \omega \|h_{T_0}\|_1 + (1 - \omega) \|h_{T_0 \cup \tilde{T} \setminus \tilde{T}_\alpha}\|_1 + 2 \left( \omega \|x_{T_0^c}\|_1 + (1 - \omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right) \right].$$

We partition  $h_{-\max(dk)}$  into two parts, i.e.,  $h_{-\max(dk)} = h^{(1)} + h^{(2)}$ , where  $h^{(1)}(i)$  equals to  $h_{-\max(dk)}(i)$  if  $|h_{-\max(dk)}(i)| > \frac{r}{t-d}$  and 0 else,  $h^{(2)}(i)$  equals to  $h_{-\max(dk)}(i)$  if  $|h_{-\max(dk)}(i)| \leq \frac{r}{t-d}$  and 0 else.

In view of the above definitions and (24),

$$\|h^{(1)}\|_1 \leq \|h_{-\max(dk)}\|_1 \leq kr.$$

Let

$$\|h^{(1)}\|_0 = m.$$

From the definition of  $h^{(1)}$ , it is clear that

$$kr \geq \|h^{(1)}\|_1 = \sum_{i \in \text{supp}(h^{(1)})} |h^{(1)}(i)| \geq \sum_{i \in \text{supp}(h^{(1)})} \frac{r}{t-d} = \frac{mr}{t-d}.$$

Namely  $m \leq k(t-d)$ . Moreover,  $\|h_{\max(dk)} + h^{(1)}\|_0 = dk + m \leq dk + k(t-d) = tk$ , and

$$\begin{aligned} \|h^{(2)}\|_1 &= \|h_{-\max(dk)}\|_1 - \|h^{(1)}\|_1 \leq kr - \frac{mr}{t-d} \\ &= (k(t-d) - m) \cdot \frac{r}{t-d}, \\ \|h^{(2)}\|_\infty &\leq \frac{r}{t-d}. \end{aligned} \quad (25)$$

By the definition of  $\delta_k$  and the fact that

$$\|Ah\|_2 \leq \|A\hat{x}^{l_2} - Ax\|_2 \leq \|y - A\hat{x}^{l_2}\|_2 + \|Ax - y\|_2 \leq 2\varepsilon, \quad (26)$$

we obtain

$$\begin{aligned}\langle A(h_{\max}(dk) + h^{(1)}), Ah \rangle &\leq \|A(h_{\max}(dk) + h^{(1)})\|_2 \|Ah\|_2 \\ &\leq \sqrt{1 + \delta_{tk}} \|h_{\max}(dk) + h^{(1)}\|_2 \cdot (2\varepsilon).\end{aligned}\quad (27)$$

Thus, using Lemma 2.1 and (25), we have  $h^{(2)} = \sum_{i=1}^N \lambda_i u_i$ ,  $\text{supp}(u_i) \subseteq \text{supp}(h^{(2)})$ ,  $\|u_i\|_1 = \|h^{(2)}\|_1$  and  $\|u_i\|_\infty \leq \frac{r}{t-d}$ , where  $u_i$  is  $(k(t-d) - m)$ -sparse, namely,  $|\text{supp}(u_i)| = \|u_i\|_0 \leq k(t-d) - m$ . Thus,

$$\begin{aligned}\|u_i\|_2 &\leq \sqrt{\|u_i\|_0} \|u_i\|_\infty \leq \sqrt{k(t-d) - m} \|u_i\|_\infty \\ &\leq \sqrt{k(t-d)} \cdot \frac{r}{t-d} \leq \sqrt{\frac{k}{t-d}} r.\end{aligned}$$

Take  $\beta_i = h_{\max}(dk) + h^{(1)} + \mu u_i$ , where  $0 \leq \mu \leq 1$ . We observe that

$$\begin{aligned}\sum_{j=1}^N \lambda_j \beta_j - \frac{1}{2} \beta_i &= h_{\max}(dk) + h^{(1)} + \mu h^{(2)} - \frac{1}{2} \beta_i \\ &= \left(\frac{1}{2} - \mu\right)(h_{\max}(dk) + h^{(1)}) - \frac{1}{2} \mu u_i + \mu h.\end{aligned}\quad (28)$$

Because  $h_{\max}(dk)$  is  $dk$ -sparse,  $h^{(1)}$  is  $m$ -sparse, and  $u_i$  is  $k(t-d) - m$ -sparse,  $\beta_i$  and  $\sum_{j=1}^N \lambda_j \beta_j - \frac{1}{2} \beta_i - \mu h$  are  $tk$ -sparse. Let

$$\begin{aligned}X &= \|h_{\max}(dk) + h^{(1)}\|_2, \\ \gamma &= \omega + (1 - \omega) \sqrt{1 + \rho - 2\alpha\rho}, \\ P &= \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1 - \omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k} \gamma}.\end{aligned}$$

Due to  $|T_0 \cup \tilde{T} \setminus \tilde{T}_\alpha| = (1 + \rho - 2\alpha\rho)k$ ,

$$\begin{aligned}
\|u_i\|_2 &\leq \sqrt{\frac{k}{t-d}} r \\
&= \sqrt{\frac{k}{t-d}} \cdot \frac{1}{k} \left[ \omega \|h_{T_0}\|_1 + (1-\omega) \|h_{T_0 \cup \tilde{T} \setminus \tilde{T}_\alpha}\|_1 + 2 \left( \omega \|x_{T_0^c}\|_1 + (1-\omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right) \right] \\
&= \frac{\omega \|h_{T_0}\|_1 + (1-\omega) \|h_{T_0 \cup \tilde{T} \setminus \tilde{T}_\alpha}\|_1}{\sqrt{k(t-d)}} + \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1-\omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k(t-d)}} \\
&\leq \frac{\omega \sqrt{k} \|h_{T_0}\|_2 + (1-\omega) \sqrt{(1+\rho-2\alpha\rho)k} \|h_{T_0 \cup \tilde{T} \setminus \tilde{T}_\alpha}\|_2}{\sqrt{k(t-d)}} + \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1-\omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k(t-d)}} \\
&\leq \frac{\omega \|h_{\max(dk)}\|_2 + (1-\omega) \sqrt{1+\rho-2\alpha\rho} \|h_{\max(dk)}\|_2}{\sqrt{t-d}} + \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1-\omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k(t-d)}} \\
&= \frac{(\omega + (1-\omega) \sqrt{1+\rho-2\alpha\rho}) \|h_{\max(dk)}\|_2}{\sqrt{t-d}} + \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1-\omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k(t-d)}} \\
&\leq \frac{\gamma}{\sqrt{t-d}} \|h_{\max(dk)} + h^{(1)}\|_2 + \frac{2 \left( \omega \|x_{T_0^c}\|_1 + (1-\omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right)}{\sqrt{k(t-d)}} \\
&= \frac{\gamma}{\sqrt{t-d}} (X + P). \tag{29}
\end{aligned}$$

We use the following identity (see (25) in [6])

$$\sum_{i=1}^N \lambda_i \left\| A \left( \sum_{j=1}^N \lambda_j \beta_j - \frac{1}{2} \beta_i \right) \right\|_2^2 = \sum_{i=1}^N \frac{\lambda_i}{4} \|A \beta_i\|_2^2. \tag{30}$$

Combining (27) and (28), we can estimate the left hand side of (30)

$$\begin{aligned}
& \sum_{i=1}^N \lambda_i \left\| A \left( \sum_{j=1}^N \lambda_j \beta_j - \frac{1}{2} \beta_i \right) \right\|_2^2 \\
&= \sum_{i=1}^N \lambda_i \left\| A \left[ \left( \frac{1}{2} - \mu \right) (h_{\max}(dk) + h^{(1)}) - \frac{1}{2} \mu u_i + \mu h \right] \right\|_2^2 \\
&= \sum_{i=1}^N \lambda_i \left\| A \left[ \left( \frac{1}{2} - \mu \right) (h_{\max}(dk) + h^{(1)}) - \frac{1}{2} \mu u_i \right] \right\|_2^2 \\
&\quad + 2 \left\langle A \left( \left( \frac{1}{2} - \mu \right) (h_{\max}(dk) + h^{(1)}) - \frac{1}{2} \mu h^{(2)} \right), \mu A h \right\rangle + \mu^2 \|A h\|_2^2 \\
&= \sum_{i=1}^N \lambda_i \left\| A \left[ \left( \frac{1}{2} - \mu \right) (h_{\max}(dk) + h^{(1)}) - \frac{1}{2} \mu u_i \right] \right\|_2^2 \\
&\quad + \mu(1 - \mu) \langle A(h_{\max}(dk) + h^{(1)}), A h \rangle \\
&\leq (1 + \delta_{tk}) \sum_{i=1}^N \lambda_i \left\| \left( \frac{1}{2} - \mu \right) (h_{\max}(dk) + h^{(1)}) - \frac{1}{2} \mu u_i \right\|_2^2 \\
&\quad + \mu(1 - \mu) \sqrt{1 + \delta_{tk}} \|h_{\max}(dk) + h^{(1)}\|_2 \cdot (2\varepsilon) \\
&= (1 + \delta_{tk}) \sum_{i=1}^N \lambda_i \left[ \left( \frac{1}{2} - \mu \right)^2 \|h_{\max}(dk) + h^{(1)}\|_2^2 + \frac{\mu^2}{4} \|u_i\|_2^2 \right] \\
&\quad + \mu(1 - \mu) \sqrt{1 + \delta_{tk}} \|h_{\max}(dk) + h^{(1)}\|_2 \cdot (2\varepsilon).
\end{aligned}$$

On the other hand, in view of the expression of  $\beta_i$ ,

$$\begin{aligned}
\sum_{i=1}^N \frac{\lambda_i}{4} \|A \beta_i\|_2^2 &= \sum_{i=1}^N \frac{\lambda_i}{4} \|A(h_{\max}(dk) + h^{(1)} + \mu u_i)\|_2^2 \\
&\geq \sum_{i=1}^N \frac{\lambda_i}{4} (1 - \delta_{tk}) \|h_{\max}(dk) + h^{(1)} + \mu u_i\|_2^2 \\
&= (1 - \delta_{tk}) \sum_{i=1}^N \frac{\lambda_i}{4} \left( \|h_{\max}(dk) + h^{(1)}\|_2^2 + \mu^2 \|u_i\|_2^2 \right).
\end{aligned}$$

It follows from the above two inequalities and (29) that

$$\begin{aligned}
0 &= \sum_{i=1}^N \lambda_i \left\| A \left( \sum_{j=1}^N \lambda_j \beta_j - \frac{1}{2} \beta_i \right) \right\|_2^2 - \sum_{i=1}^N \frac{\lambda_i}{4} \|A \beta_i\|_2^2 \\
&\leq (1 + \delta_{tk}) \sum_{i=1}^N \lambda_i \left[ \left( \frac{1}{2} - \mu \right)^2 \|h_{\max}(dk) + h^{(1)}\|_2^2 + \frac{\mu^2}{4} \|u_i\|_2^2 \right] \\
&\quad + \mu(1 - \mu) \sqrt{1 + \delta_{tk}} \|h_{\max}(dk) + h^{(1)}\|_2 \cdot (2\varepsilon) \\
&\quad - (1 - \delta_{tk}) \sum_{i=1}^N \frac{\lambda_i}{4} \left( \|h_{\max}(dk) + h^{(1)}\|_2^2 + \mu^2 \|u_i\|_2^2 \right) \\
&= \sum_{i=1}^N \lambda_i \left\{ \left( (1 + \delta_{tk}) \left( \frac{1}{2} - \mu \right)^2 - \frac{1}{4} (1 - \delta_{tk}) \right) \right. \\
&\quad \cdot \left. \|h_{\max}(dk) + h^{(1)}\|_2^2 + \frac{1}{2} \delta_{tk} \mu^2 \|u_i\|_2^2 \right\} \\
&\quad + \mu(1 - \mu) \sqrt{1 + \delta_{tk}} \|h_{\max}(dk) + h^{(1)}\|_2 \cdot (2\varepsilon) \\
&\leq \left[ (1 + \delta_{tk}) \left( \frac{1}{2} - \mu \right)^2 - \frac{1}{4} (1 - \delta_{tk}) + \frac{\delta_{tk} \mu^2 \gamma^2}{2(t-d)} \right] X^2 \\
&\quad + \left[ \mu(1 - \mu) \sqrt{1 + \delta_{tk}} \cdot (2\varepsilon) + \frac{\delta_{tk} \mu^2 \gamma^2 P}{t-d} \right] X + \frac{\delta_{tk} \mu^2 \gamma^2 P^2}{2(t-d)} \\
&= \left[ (\mu^2 - \mu) + \left( \frac{1}{2} - \mu + \left( 1 + \frac{\gamma^2}{2(t-d)} \right) \mu^2 \right) \delta_{tk} \right] X^2 \\
&\quad + \left[ \mu(1 - \mu) \sqrt{1 + \delta_{tk}} \cdot (2\varepsilon) + \frac{\delta_{tk} \mu^2 \gamma^2 P}{t-d} \right] X + \frac{\delta_{tk} \mu^2 \gamma^2 P^2}{2(t-d)}.
\end{aligned}$$

Taking  $\mu = \frac{\sqrt{(t-d)(t-d+\gamma^2)} - (t-d)}{\gamma^2}$ , we obtain

$$\begin{aligned}
&-\frac{t-d+\gamma^2}{t-d} \mu^2 \left( \sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk} \right) X^2 \\
&+ \left( \frac{t-d+\gamma^2}{t-d} \mu^2 \sqrt{\frac{t-d}{t-d+\gamma^2}} \sqrt{1 + \delta_{tk}} \cdot (2\varepsilon) + \frac{\delta_{tk} \mu^2 \gamma^2 P}{t-d} \right) X + \frac{\delta_{tk} \mu^2 \gamma^2 P^2}{2(t-d)} \geq 0.
\end{aligned}$$

Namely,

$$\begin{aligned}
&\frac{\mu^2}{t-d} \left[ -(t-d+\gamma^2) \left( \sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk} \right) X^2 \right. \\
&\quad + \left( \sqrt{(t-d)(t-d+\gamma^2)} (1 + \delta_{tk}) \cdot (2\varepsilon) \right. \\
&\quad \left. \left. + \delta_{tk} \gamma^2 P \right) X + \frac{\delta_{tk} \gamma^2 P^2}{2} \right] \geq 0,
\end{aligned}$$



which is a second-order inequality for  $X$ . Hence, we have

$$\begin{aligned}
X &\leq \left\{ \left( 2\varepsilon \sqrt{(t-d)(t-d+\gamma^2)(1+\delta_{tk})} + \delta_{tk}\gamma^2 P \right) \right. \\
&\quad + \left[ \left( 2\varepsilon \sqrt{(t-d)(t-d+\gamma^2)(1+\delta_{tk})} + \delta_{tk}\gamma^2 P \right)^2 \right. \\
&\quad \left. \left. + 2(t-d+\gamma^2) \left( \sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk} \right) \delta_{tk}\gamma^2 P^2 \right]^{1/2} \right\} \\
&\quad \cdot \left( 2(t-d+\gamma^2) \sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk} \right)^{-1} \\
&\leq \frac{\sqrt{(t-d)(t-d+\gamma^2)(1+\delta_{tk})}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})} (2\varepsilon) \\
&\quad + \frac{2\delta_{tk}\gamma^2 + \sqrt{2(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})\delta_{tk}\gamma^2}}{2(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})} P.
\end{aligned}$$

From (23) and the representation of  $P$ , it is clear that

$$\begin{aligned}
\|h_{-\max(dk)}\|_1 &\leq \|h_{\max(dk)}\|_1 + 2 \left( \omega \|x_{T_0^c}\|_1 + (1-\omega) \|x_{\tilde{T}^c \cap T_0^c}\|_1 \right) \\
&= \|h_{\max(dk)}\|_1 + P\sqrt{k}\gamma.
\end{aligned}$$

It follows from Lemma 2.2 that

$$\|h_{-\max(dk)}\|_2 \leq \|h_{\max(dk)}\|_2 + \frac{P\gamma}{\sqrt{d}}.$$

Thus, we have the estimate of  $\|h\|_2$  by the above inequalities

$$\begin{aligned}
\|h\|_2 &= \sqrt{\|h_{\max(dk)}\|_2^2 + \|h_{-\max(dk)}\|_2^2} \\
&\leq \sqrt{\|h_{\max(dk)}\|_2^2 + \left(\|h_{\max(dk)}\|_2 + \frac{P\gamma}{\sqrt{d}}\right)^2} \\
&\leq \sqrt{2}\|h_{\max(dk)}\|_2 + \frac{P\gamma}{\sqrt{d}} \\
&\leq \sqrt{2}\|h_{\max(dk)} + h^{(1)}\|_2 + \frac{P\gamma}{\sqrt{d}} \\
&= \sqrt{2}X + \frac{P\gamma}{\sqrt{d}} \\
&\leq \frac{\sqrt{2(t-d)(t-d+\gamma^2)(1+\delta_{tk})}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})}(2\varepsilon) \\
&\quad + \left( \frac{\sqrt{2}\delta_{tk}\gamma^2 + \sqrt{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})\delta_{tk}\gamma^2}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})} + \frac{\gamma}{\sqrt{d}} \right) P \\
&= \frac{\sqrt{2(t-d)(t-d+\gamma^2)(1+\delta_{tk})}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})}(2\varepsilon) \\
&\quad + \left( \frac{\sqrt{2}\delta_{tk}\gamma^2 + \sqrt{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})\delta_{tk}\gamma^2}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})} + \frac{\gamma}{\sqrt{d}} \right) \frac{2\left(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1\right)}{\sqrt{k}\gamma} \\
&= \frac{\sqrt{2(t-d)(t-d+\gamma^2)(1+\delta_{tk})}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})}(2\varepsilon) \\
&\quad + \left( \frac{\sqrt{2}\delta_{tk}\gamma + \sqrt{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})\delta_{tk}}}{(t-d+\gamma^2)(\sqrt{\frac{t-d}{t-d+\gamma^2}} - \delta_{tk})} + \frac{1}{\sqrt{d}} \right) \frac{2\left(\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_1\right)}{\sqrt{k}}.
\end{aligned}$$

If  $tk$  is not an integer, taking  $t' = \lceil tk \rceil / k$ , then  $t'k$  is an integer and  $t < t'$ . Thus we have

$$\delta_{t'k} = \delta_{tk} < \sqrt{\frac{t-d}{t-d+\gamma^2}} < \sqrt{\frac{t'-d}{t'-d+\gamma^2}}.$$

Then we can prove the result the same as the proof above by working on  $\delta_{t'k}$ . So, we obtain (19).

Next, we prove (21). The proof of (21) is similar to the proof of (19). We only need to replace (26) and (27) with the following (31) and (32), respectively. We also can get (21).

$$\begin{aligned}
\|A^T Ah\|_\infty &= \|A^T A(\widehat{x}^{DS} - x)\|_\infty \\
&\leq \|A^T (A\widehat{x}^{DS} - y)\|_\infty + \|A^T (y - Ax)\|_\infty \\
&\leq 2\varepsilon,
\end{aligned} \tag{31}$$

$$\begin{aligned}
\langle A(h_{\max}(dk) + h^{(1)}), Ah \rangle &= \langle h_{\max}(dk) + h^{(1)}, A^T Ah \rangle \\
&\leq \|h_{\max}(dk) + h^{(1)}\|_1 \|A^T Ah\|_\infty \\
&\leq \sqrt{tk} \|h_{\max}(dk) + h^{(1)}\|_2 \cdot (2\varepsilon).
\end{aligned} \tag{32}$$

This completes the proof of Theorem 3.1.  $\square$

*Proof of Theorem 3.2.* For  $d = 1$ , we have  $\gamma = \omega + (1 - \omega)\sqrt{1 + \rho - 2\alpha\rho} \leq 1$ . Moreover, for all  $\varepsilon > 0$  and  $k \geq \frac{6}{\varepsilon}$ , we define

$$m' = \frac{1 + \sqrt{1 - \gamma^2}}{\gamma^2} \left( t - 1 + \sqrt{(t - 1)(t - 1 + \gamma^2)} \right) k$$

and  $N \geq k + m'$ . Since  $t \geq 1 + \frac{(1 - \sqrt{1 - \gamma^2})^2}{\gamma^2 + 2(1 - \sqrt{1 - \gamma^2})}$ , we obtain  $m' \geq k$ . Let  $m$  be the largest integer strictly smaller than  $m'$ , then  $m < m'$  and  $m' - m \leq 1$ . We take

$$x_1 = \frac{1}{\sqrt{k + \frac{mk^2}{m'^2}}} \left( \underbrace{1, \dots, 1}_{k - \alpha\rho k}, \underbrace{-\frac{k}{m'}, \dots, -\frac{k}{m'}}_{\rho k}, \underbrace{1, \dots, 1}_{\alpha\rho k}, \underbrace{-\frac{k}{m'}, \dots, -\frac{k}{m'}}_{m - \rho k}, 0, \dots, 0 \right) \in \mathbb{R}^N,$$

if  $m > \rho k$ ; or take

$$x_1 = \frac{1}{\sqrt{k + \frac{mk^2}{m'^2}}} \left( \underbrace{1, \dots, 1}_{k - \alpha\rho k}, \underbrace{-\frac{k}{m'}, \dots, -\frac{k}{m'}}_{\rho k}, \underbrace{0, \dots, 0, 1, \dots, 1}_{\alpha\rho k}, 0, \dots, 0 \right) \in \mathbb{R}^N,$$

if  $m \leq \rho k$ . It is easy to know  $\|x_1\|_2 = 1$ . Define the linear map  $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$  by

$$\begin{aligned}
Ax &= \sqrt{1 + \sqrt{\frac{t - d}{t - d + \gamma^2}}} (x - \langle x_1, x \rangle x_1) \\
&= \sqrt{1 + \sqrt{\frac{t - 1}{t - 1 + \gamma^2}}} (x - \langle x_1, x \rangle x_1),
\end{aligned}$$

for all  $x \in \mathbb{R}^N$ . Then for any  $\lceil tk \rceil$ -sparse vector  $x$ , we get

$$\|Ax\|_2^2 = \left( 1 + \sqrt{\frac{t - 1}{t - 1 + \gamma^2}} \right) (\|x\|_2^2 - |\langle x_1, x \rangle|^2).$$

Hence, using Cauchy-Schwarz inequality and the fact that  $m' \geq k$ ,  $m' - m \leq 1$  and

$$\frac{m'^2 + k^2(t-1)}{m'^2 + m'k} = \frac{2\sqrt{t-1}(\sqrt{t-1+\gamma^2} - \sqrt{t-1})}{\gamma^2},$$

we have

$$\begin{aligned} 0 &\leq |\langle x_1, x \rangle|^2 \leq \|x\|_2^2 \cdot \sum_{i \in \text{supp}(x)} |x_1(i)|^2 \\ &\leq \|x\|_2^2 \cdot \|x_{1, \max(\lceil tk \rceil)}\|_2^2 \\ &= \|x\|_2^2 \cdot \frac{m'^2 + k(\lceil tk \rceil - k)}{m'^2 + m'k} \\ &\leq \frac{m'^2 + k^2(t-1) + k}{m'^2 + m'k} \|x\|_2^2 \\ &= \frac{m'^2 + k^2(t-1) + k}{m'^2 + m'k} \cdot \frac{m'^2 + m'k}{m'^2 + m'k} \|x\|_2^2 \\ &= \frac{m'^2 + k^2(t-1) + k}{m'^2 + m'k} \cdot \frac{1}{1 - \frac{k(m'-m)}{m'^2 + m'k}} \|x\|_2^2 \\ &= \frac{m'^2 + k^2(t-1)}{m'^2 + m'k} \cdot \frac{m'^2 + k^2(t-1) + k}{m'^2 + k^2(t-1)} \cdot \frac{1}{1 - \frac{k(m'-m)}{m'^2 + m'k}} \|x\|_2^2 \\ &\leq \frac{2\sqrt{t-1}(\sqrt{t-1+\gamma^2} - \sqrt{t-1})}{\gamma^2} \cdot \left(1 + \frac{1}{tk}\right) \cdot \frac{1}{1 - \frac{1}{2k}} \|x\|_2^2 \\ &\leq \frac{2\sqrt{t-1}(\sqrt{t-1+\gamma^2} - \sqrt{t-1})}{\gamma^2} \cdot \left(1 + \frac{3}{k}\right) \|x\|_2^2 \\ &\leq \left( \frac{2\sqrt{t-1}(\sqrt{t-1+\gamma^2} - \sqrt{t-1})}{\gamma^2} + \frac{3}{k} \right) \|x\|_2^2. \end{aligned}$$

Consequently,

$$\begin{aligned}
& \left(1 + \sqrt{\frac{t-1}{t-1+\gamma^2}} + \varepsilon\right) \|x\|_2^2 \\
& \geq \left(1 + \sqrt{\frac{t-1}{t-1+\gamma^2}}\right) \|x\|_2^2 \geq \|Ax\|_2^2 \\
& \geq \left(1 + \sqrt{\frac{t-1}{t-1+\gamma^2}}\right) \cdot \left(1 - \frac{2\sqrt{t-1}(\sqrt{t-1+\gamma^2} - \sqrt{t-1})}{\gamma^2} - \frac{3}{k}\right) \|x\|_2^2 \\
& = \left[ \left(1 + \sqrt{\frac{t-1}{t-1+\gamma^2}}\right) \cdot \left(1 - \frac{2\sqrt{t-1}(\sqrt{t-1+\gamma^2} - \sqrt{t-1})}{\gamma^2}\right) \right. \\
& \quad \left. - \left(1 + \sqrt{\frac{t-1}{t-1+\gamma^2}}\right) \frac{3}{k} \right] \|x\|_2^2 \\
& = \left[ 1 - \sqrt{\frac{t-1}{t-1+\gamma^2}} - \left(1 + \sqrt{\frac{t-1}{t-1+\gamma^2}}\right) \frac{3}{k} \right] \|x\|_2^2 \\
& \geq \left(1 - \sqrt{\frac{t-1}{t-1+\gamma^2}} - \varepsilon\right) \|x\|_2^2,
\end{aligned}$$

which deduces  $\delta_{tk} \leq \sqrt{\frac{t-1}{t-1+\gamma^2}} + \varepsilon$ . Next, we define

$$\begin{aligned}
x_0 &= (\overbrace{1, \dots, 1}^{k-\alpha\rho k}, \overbrace{0, \dots, 0}^{\rho k}, \overbrace{1, \dots, 1}^{\alpha\rho k}, 0, \dots, 0) \in \mathbb{R}^N, \\
\eta_0 &= (\overbrace{0, \dots, 0}^{k-\alpha\rho k}, \overbrace{\frac{k}{m'}, \dots, \frac{k}{m'}}_{\rho k}, \overbrace{0, \dots, 0}^{\alpha\rho k}, \overbrace{\frac{k}{m'}, \dots, \frac{k}{m'}}_{m-\rho k}, 0, \dots, 0) \in \mathbb{R}^N, \text{ if } m > \rho k, \\
\text{or } \eta_0 &= (\overbrace{0, \dots, 0}^{k-\alpha\rho k}, \overbrace{\frac{k}{m'}, \dots, \frac{k}{m'}}_m, \overbrace{0, \dots, 0}^{\rho k}, \overbrace{0, \dots, 0}^{\alpha\rho k}, 0, \dots, 0) \in \mathbb{R}^N, \text{ if } m \leq \rho k,
\end{aligned}$$

where  $\|x_0\|_{1,w} = k$ ,  $\|\eta_0\|_{1,w} \leq m \cdot \frac{k}{m'} < k$ . Obviously,  $x_0$  is  $k$ -sparse,  $x_1 = \frac{1}{\sqrt{k + \frac{mk^2}{m'^2}}}(x_0 - \eta_0)$  and  $\|\eta_0\|_{1,w} < \|x_0\|_{1,w}$ . In view of  $Ax_1 = 0$ , we have  $Ax_0 = A\eta_0$ .

Thus, in the noiseless case  $y = Ax_0$ , suppose that the weighted  $l_1$  minimization method (7) can exactly recover  $x_0$ , i.e.,  $\hat{x} = x_0$ . According to the definition of  $\hat{x}$  and  $y = A\eta_0$ , it contradicts that  $\|\eta_0\|_{1,w} < \|x_0\|_{1,w} = \|\hat{x}\|_{1,w}$ .

In the noise case  $y = Ax_0 + z$ , suppose that the weighted  $l_1$  minimization method (7) can stably recover  $x_0$ , i.e.,  $\lim_{z \rightarrow 0} \hat{x} = x_0$ . We observe that  $y - A(\hat{x} - x_0 + \eta_0) = y - A\hat{x} \in \mathcal{B}$ , thus  $\|\hat{x}\|_{1,w} \leq \|\hat{x} - x_0 + \eta_0\|_{1,w}$ . As  $z \rightarrow 0$ ,  $\|x_0\|_{1,w} \leq \|\eta_0\|_{1,w}$ . It contradicts that  $\|\eta_0\|_{1,w} < \|x_0\|_{1,w}$ .

Hence, the weighted  $l_1$  minimization method (7) fails to exactly and stably recover  $x_0$  based on  $y$  and  $A$ .  $\square$

## References

- [1] R. Baraniuk and P. Steeghs, Compressive radar imaging, in Proc. IEEE Radar Conf., 2007, pp. 128-133.
- [2] R. V. Borries, C. Miosso and C. Potes, Compressed sensing using prior information, in 2nd IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMPSAP 2007, 12-14, 2007, pp. 121-124.
- [3] T. T. Cai, L. Wang and G. W. Xu, Shifting inequality and recovery of sparse signals, IEEE Trans. Signal Process, 58(3), pp. 1300-1308, 2010.
- [4] T. T. Cai, L. Wang and G. W. Xu, New bounds for restricted isometry constants, IEEE Trans. Inform. Theory, 56(9), pp. 4388-4394, 2010.
- [5] T. T. Cai, G. W. Xu, J. Zhang, On recovery of sparse signal via  $l_1$  minimization, IEEE Trans. Inf. Theory, 55(7), pp. 3388-3397, 2009.
- [6] T. T. Cai and A. Zhang, Spares representation of a polytope and recovery of sparse signals and low-rank matrices, IEEE Trans. Inform. Theory, 60(1), pp. 122-132, 2014.
- [7] T. T. Cai and A. Zhang, Compressed sensing and affine rank minimization under restricted isometry, IEEE Trans. Signal Process., 61(13), pp. 3279-3290, 2013.
- [8] T. T. Cai and A. Zhang, Sharp RIP bound for sparse signal and low-rank matrix recovery, Appl. Comput. Harmon. Anal., 35, pp. 74-93, 2013.
- [9] E. J. Candès, J. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Commum. Pure Appl. Math., 59, pp. 1207-1223, 2006.
- [10] E. Candès and T. Tao, Decoding by linear programming, IEEE Trans. Inf. Theory, 51(12), pp. 4203-4215, 2005.
- [11] Q. Du and J. E. Fowler, Hyperspectral image compression using jpeg2000 and principal component analysis, IEEE Geosci. Remote Sens. Lett., 4(4), pp. 201-205, 2007.
- [12] M. P. Friedlander, H. Mansour, R. Saab and O. Yilmaz, Recovering compressively sampled signals using partial support information, IEEE Transactions Information Theory, 58(2), pp. 1122-1134, 2012.
- [13] M. Herman and T. Strohmer, High-resolution radar via compressed sensing, IEEE Signal Process. Mag., 57(6), pp. 2275-2284, 2009.

- [14] L. Jacques, A short note compressed sensing with partially known signal support, *Signal Process.*, 90, pp. 3308-3312, 2010.
- [15] M. A. Khajehnejad, W. Xu, A. S. Avestimehr and B. Hassibi, Weighted  $l_1$  minimization for sparse recovery with prior information, in *IEEE Int. Symp. Information Theory, ISIT 2009*, 2009, pp. 483-487.
- [16] W. Lu and N. Vaswani, Exact reconstruction conditions and error bounds for regularized modified basis pursuit, in *Proc. Asilomar Conf. on Signals, Systems and Computers*, 2010.
- [17] W. Lu and N. Vaswani, Modified basis pursuit denoising (modifiedbpdn) for noisy compressive sensing with partially known signal support, in *IEEE Int. Conf. Acoustics Speech and Signal Processing (ICASSP)*, 2010, 14-19, 2010, pp. 3926-3929.
- [18] M. Lustig, J. M. Santos, J. H. Lee, D. L. Donoho and J. M. Pauly, Application of compressed sensing for rapid MR imaging, in *SPARS*, (Rennes, France), 2005.
- [19] Q. Mo and S. Li, New bounds on the restricted isometry constant  $\delta_{2k}$ , *Appl. Comput. Harmon. Anal.*, 31(3), pp. 3335460-468, 2011.
- [20] F. Parvaresh, H. Vikalo, S. Misra and B. Hassibi, Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays, *IEEE J. Sel. Top. Signal Process.*, 2(3), pp. 275-285, 2008.
- [21] N. Vaswani and W. Lu, Modified-CS: Modifying compressive sensing for problems with partially known support, *IEEE Trans. Signal Process.*, 58(9), pp. 4595-4607, 2010.
- [22] N. Vaswani and W. Lu, Modified-CS: Modifying compressive sensing for problems with partially known support, *IEEE Int. Symp. Information Theory, ISIT 2009*, pp. 488-492, 2009.
- [23] J. Zhang, D. Zhu and G. Zhang, Adaptive compressed sensing radar oriented toward cognitive decetion in dynamic sparse target scene, *IEEE Trans. Signal Process.*, 60(4), pp. 1718-1729, 2012.